Multiple Sclerosis and Related Disorders 29 (2019) 55–61

Contents lists available at ScienceDirect

Multiple Sclerosis and Related Disorders

journal homepage: www.elsevier.com/locate/msard

Systematic review and network meta-analysis comparing ocrelizumab with other treatments for relapsing multiple sclerosis

Rachael McCool[a], Katy Wilson[a], Mick Arber[a], Kelly Fleetwood[b], Sydney Toupin[b], Howard Thom[c], Iain Bennett[d], Susan Edwards[d,*]

[a] York Health Economics Consortium, Enterprise House, Innovation Way, University of York, York YO10 5NQ, UK
[b] Quantics Biostatistics, West End House, 28 Drumsheugh Gardens, Edinburgh EH3 7RN, UK
[c] Bristol Medical School: Population Health Sciences, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK
[d] Health Economics and Evidence Synthesis, Global Access Center of Excellence, F. Hoffmann-La Roche Ltd, Grenzacherstrasse 124, Basel, CH-4070, Switzerland

ARTICLE INFO

Keywords:
Relapsing multiple sclerosis
Disease-modifying therapy
Ocrelizumab
Systematic literature review
Network meta-analysis
Anti-CD20

ABSTRACT

Background: Ocrelizumab was approved for the treatment of relapsing multiple sclerosis (RMS) and primary progressive multiple sclerosis (PPMS) by the US Food and Drug Administration in March 2017 and by the European Medicines Agency in January 2018. These approvals were based on two pivotal randomized controlled trials (RCTs), OPERA I and OPERA II, comparing ocrelizumab 600 mg with an active comparator, interferon β-1a 44 μg (Rebif), and the first trial with positive results in patients with PPMS, which compared ocrelizumab with placebo. However, direct evidence of the efficacy and safety of ocrelizumab in RMS compared with other disease-modifying therapies (DMTs) approved for RMS is not available from RCTs. In the absence of such RCTs, network meta-analyses (NMAs) were conducted to compare indirectly the relative efficacy and safety of ocrelizumab with all other approved DMTs for the treatment of RMS.

Methods: Systematic literature searches were conducted in MEDLINE, Embase, the Cochrane Library, trial registers, relevant conference websites and health technology assessment agency websites. Eligible RCTs evaluated approved treatments for multiple sclerosis (MS) in which more than 75% of patients had a relapsing form of MS. NMAs were conducted for four efficacy and three safety outcomes, and treatment hierarchies were generated for each outcome using surface under the cumulative ranking curve (SUCRA) values.

Results: Results suggest that ocrelizumab has superior efficacy to 10 of the 17 treatments in the 12-week confirmed disability progression network and 12 of the 17 treatments in the annualized relapse rate network (both including placebo). The efficacy of ocrelizumab was comparable with the other treatments in both networks. In the serious adverse events and discontinuation due to adverse events networks, ocrelizumab demonstrated a safety profile comparable with all other treatments (including placebo). SUCRA values consistently ranked ocrelizumab among the most effective or tolerable treatments across all outcomes.

Conclusions: Results suggest that ocrelizumab has an efficacy superior to or comparable with all other currently approved DMTs across all endpoints analyzed, and a similar safety profile, indicating it offers a valuable package for the treatment of patients with RMS.

1. Introduction

Multiple sclerosis (MS) is a chronic autoimmune disease, characterized by inflammation of the central nervous system that leads to progressive neuro-axonal degeneration (Lavery et al., 2014). There is no cure for MS, but 14 disease-modifying therapies (DMTs; following the withdrawal of daclizumab from the market on 30 April 2018, it is currently 13) have been approved by the European Medicines Agency

Abbreviations: AE, adverse event; ARR, annualized relapse rate; CDP, confirmed disability progression; DIC, deviance information criterion; DMT, disease-modifying therapy; DSU TSD, Decision Support Unit Technical Support Document; EMA, European Medicines Agency; ICER, Institute for Clinical and Economic Review; ITT, intent-to-treat; JAGS, Just Another Gibbs Sampler; MS, multiple sclerosis; NEDA, no evidence of disease activity; NICE, The National Institute for Health and Care Excellence; NMA, network meta-analysis; PPMS, primary progressive multiple sclerosis; RCT, randomized controlled trial; RMS, relapsing multiple sclerosis; RRMS, relapsing-remitting multiple sclerosis; SAE, serious adverse event; SLR, systematic literature review; SUCRA, surface under the cumulative ranking curve

* Corresponding author.
E-mail address: susan.edwards.se1@roche.com (S. Edwards).

https://doi.org/10.1016/j.msard.2018.12.040
Received 15 June 2018; Received in revised form 5 December 2018; Accepted 31 December 2018

(EMA) for the treatment of patients with relapsing forms of MS (RMS). These DMTs aim to slow progression of disability, reduce the number and severity of relapses, and diminish the impact of MS on health-related quality of life (Torkildsen et al., 2016). Selection of treatment requires a compromise between the efficacy and safety profiles, which vary across treatments. Given that DMTs with higher efficacy are often associated with increased risk of serious adverse events (SAEs), they are usually reserved for use later in the disease course (Gajofatto, 2015). However, recent evidence suggests that early treatment with high-efficacy DMTs may lead to improved disease control (Merkel et al., 2017; Rotstein et al., 2015).

Comparisons of efficacy and safety across different DMTs are needed to inform treatment decisions for patients with MS. Head-to-head randomized controlled trials (RCTs) are the gold standard for generating this evidence. However, a single RCT of all DMTs has not been conducted. In the absence of such a trial, network meta-analyses (NMAs) can be used to compare treatments using both direct comparisons of interventions within clinical trials and indirect comparisons based on a common comparator, often placebo (Li et al., 2011). Several such NMAs have compared the different DMTs for RMS (Fogarty et al., 2016; Huisman et al., 2017; Institute for Clinical and Economic Review, 2017; Tolley et al., 2015; Tramacere et al., 2015; Tsivgoulis et al., 2015; Siddiqui et al., 2017; Melendez-Torres et al., 2017). However, these NMAs have used differing methodologies, have not used all available data by combining hazard ratio and count data, (Watkins, 2018) and did not consider the trade-off between efficacy and safety with surface under the cumulative ranking curve (SUCRA) values.

Ocrelizumab is a humanized anti-CD20 monoclonal antibody approved by the EMA in January 2018 for the treatment of adult patients with RMS with active disease (defined by clinical or imaging features), (European Medicines Agency (EMA), 2018) and has been received by over 70,000 patients worldwide since US Food and Drug Administration approval in March 2017 (Hoffmann-La Roche Ltd, 2018). The objective of this study was to conduct a systematic literature review (SLR) and NMA to evaluate the relative efficacy and safety of ocrelizumab compared with all approved DMTs for the treatment of patients with RMS, across seven measures of efficacy and safety.

2. Materials and methods

2.1. Identification and selection of relevant trials

An SLR was performed following National Institute of Health and Care Excellence (NICE) guidelines to identify all RCTs assessing the efficacy and safety of DMTs used for the treatment of patients with RMS (National Institute for Health and Care Excellence (NICE), 2015). Searches for relevant trials were conducted in October 2014, with updates performed in November 2016 and July 2017. (**Appendix A.1; Table D1; Table D2**). The review protocol was not registered.

Following identification of records, duplicates were removed, and titles and abstracts were screened against the eligibility criteria; full texts of the included records were then reviewed in a second review round, against the same criteria. Both rounds of screening were carried out by two independent reviewers, with disagreements resolved by a third. Eligibility criteria included trials with a duration of at least 12 weeks, evaluating suitable interventions in which more than 75% of patients had RMS (**Table D3**). Eligible interventions are listed in **Table D3**. These included both doses of cladribine (3.5 mg/kg and 5 mg/kg), which had received a positive opinion from the Committee for Medicinal Products for Human Use at the time of analyses (July 2017). Since analysis, only the 3.5 mg/kg dose has been approved by the EMA. Similarly, daclizumab was included because it was approved at the time of analyses, but has since been withdrawn by the manufacturer because of concerns surrounding its benefit-risk profile.

2.2. Data extraction and risk of bias

For trials that met the eligibility criteria, data relating to trial design and methodology, details of interventions, patient eligibility criteria and baseline characteristics, and outcome measures were extracted. Data were extracted for the intent-to-treat (ITT) populations of the trials and for the highly active and rapidly evolving severe RMS subgroups. The risk of bias in each individual trial was assessed using the Cochrane Risk of Bias Tool (**Appendix A.2; Table D4**) (Higgins et al., 2011).

2.3. Network meta-analyses

The key efficacy outcomes of interest for the NMAs were 12-week confirmed disability progression (12-week CDP) and annualized relapse rate (ARR; the primary outcome of most phase 3 clinical trials in MS, including the OPERA I and OPERA II ocrelizumab trials). The key safety outcomes of interest were SAEs and discontinuation of treatment due to an adverse event (AE). In addition, the following outcomes were also analyzed: 24-week confirmed disability progression (24-week CDP), proportion of patients who remained relapse-free and all-cause discontinuation of treatment. A feasibility assessment was conducted to assess the similarity of the trials (**Appendix A.3**) (Dias et al., 2011b). Trials which were suitably similar were included in the NMAs, providing there was sufficient similarity across the definitions of the outcomes (**Appendix A.4; Table D5**).

All outcomes were analyzed using standard Bayesian approaches as described in the NICE Decision Support Unit Technical Support Document (DSU TSD) 2 (**Appendix A.5**) (Dias et al., 2011b). The analyses used the ITT population results, where available. The base-case analyses used random effects models because these are considered more appropriate than fixed effect models when there is heterogeneity between patient populations across trials. Three types of model were applied for the base-case analysis, depending on the outcome: (i) a survival model for 12-week CDP and 24-week CDP; (ii) a Poisson model for ARR; and (iii) a binomial model for proportion of patients relapse-free and the three safety outcomes (SAEs, discontinuation due to AEs and all-cause discontinuation). For the between-trial variance, informative prior distributions were assumed following Turner et al. for both the survival and binomial models, with a vague prior distribution assumed for the Poisson models (**Appendix A.5.4**) (Turner et al., 2015). Vague prior distributions were used for all other parameters. The input data and code used for each outcome are presented in **Appendix A.5.6** and **Tables D8–D14**.

2.4. Sensitivity analyses and network meta-regressions

For each outcome, two sensitivity analyses were conducted to evaluate the assumptions of the base-case NMA: (i) a fixed effect model instead of a random effects model and (ii) a random effects model with a different prior distribution for the between-trial variance. These were vague priors for the survival and binomial models, and an alternative vague prior for the Poisson model (**Appendix A.5.4**). For each outcome, meta-regression was also used to explore whether the time at which outcomes were observed (i.e. follow-up time, on a continuous scale) influenced the relative treatment effects.

To compare the fit of the base-case models to the data with the sensitivity analyses and the meta-regression models, posterior mean residual deviance and deviance information criterion (DIC) values were calculated.

2.5. Heterogeneity and inconsistency assessments

Heterogeneity and inconsistency assessments were conducted for the base-case NMAs (**Appendix A.5.8**). The I^2 statistic was used to measure heterogeneity, with thresholds of 0–40%, 30–60%, 50–90% and 75–100% corresponding to “might not be important”, “may

represent moderate heterogeneity", "may represent substantial heterogeneity", and "considerable heterogeneity", respectively, according to the Cochrane Handbook (Higgins et al., 2011). Inconsistency was evaluated by comparing the DIC values of the base-case NMA model to an inconsistency model as recommended in the NICE DSU TSD 4 (Dias et al., 2011a).

2.6. Surface under the cumulative ranking curve (SUCRA)

The probabilities of a treatment being at each possible rank were summarized using SUCRA values (Salanti et al., 2011). To demonstrate the efficacy and safety profile of each treatment, SUCRA values for the two key efficacy outcomes (12-week CDP and ARR) and two key safety outcomes (SAE and discontinuation due to AEs) were plotted on a radar plot. SUCRA values range from 0% (the treatment is certainly ranked last) to 100% (the treatment is certainly ranked first; **Appendix A.5.9**). A similar approach was reported by Tramacere et al. (2015) although SUCRA values on only two dimensions were considered.

2.7. Analysis

All analyses were conducted using Just Another Gibbs Sampler (JAGS) version 4.1.0 or above, and R version 3.1.2 or above. The package *R2JAGS* was used to run JAGS from within R (Plummer, 2018; R Development Core Team 2018; Su and Yajima, 2018).

3. Results

After full text review of records identified in the SLR, a total of 183 records met the eligibility criteria for inclusion in the NMA, relating to 46 RCTs (Fig. 1). Of these 46, two RCTs were excluded for inappropriate treatment regimens/doses, and a further 11 RCTs with duration less than 48 weeks were excluded, as these trials were not designed to study clinical outcomes and were therefore considered too different from the other trials for inclusion in NMAs (**Appendix B.2; Table D6**). Therefore, 33 trials were included (**Table D7; Table D17**). Analysis of RMS subgroup data is not reported here because: (i) of issues with publication bias (favourable subgroups are more likely to be reported); (ii) of breaking of randomization in NMAs when subgroup data are used; and (iii) data were not reported by many of the trials, so the resulting networks were disconnected. Results for all outcomes versus placebo are presented in **Appendix B.3.1 (Table D16; Table D18**).

3.1. Key efficacy and safety outcomes

3.1.1. 12-week confirmed disability progression network

The base-case network for 12-week CDP included 17 different treatments, including placebo, from 22 trials (Table D7). Comparisons of ocrelizumab versus other DMTs provided evidence that ocrelizumab was more effective in reducing the risk of 12-week CDP than 10 other treatments, including placebo (95% credible intervals below 1; (Fig. 2, panel a; Appendix B.3.1; Table D15). The probability ocrelizumab was more effective than the six remaining treatments was greater than 50% in each case. There was no evidence to suggest any treatment was more effective than ocrelizumab.

3.1.2. Annualized relapse rate network

The base-case network for ARR included 17 different treatments, including placebo, from 30 trials (Table D7). Etemadifar et al., was excluded because it reported insufficient data on the ARR (Etemadifar et al., 2006). Comparisons of ocrelizumab versus other DMTs provided evidence that ocrelizumab was more effective in reducing the ARR than 12 other treatments, including placebo (Fig. 2, panel b; Appendix B.3.1; Table D15). The probability ocrelizumab was more effective than two of the four remaining treatments was greater than 50% in each case. There was no evidence to suggest any treatment was more effective than ocrelizumab.

3.1.3. Serious adverse events network

The base-case network for serious adverse events included 17 different treatments (including placebo) from 24 trials (**Table D17**). Comparisons of ocrelizumab versus other DMTs showed that there was no evidence of a difference between ocrelizumab and any other treatments in the risk of SAEs (Fig. 2**, panel c**).

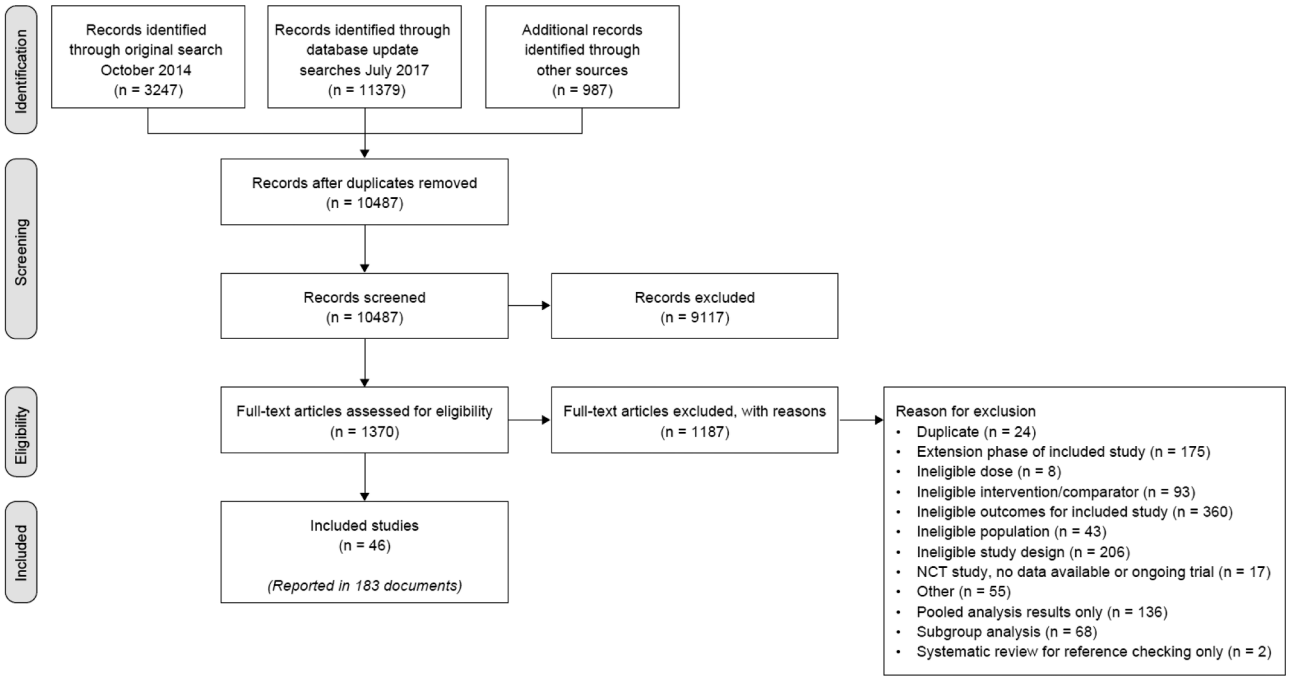

Fig. 1. Flow of information for identification of studies in the systematic literature review.

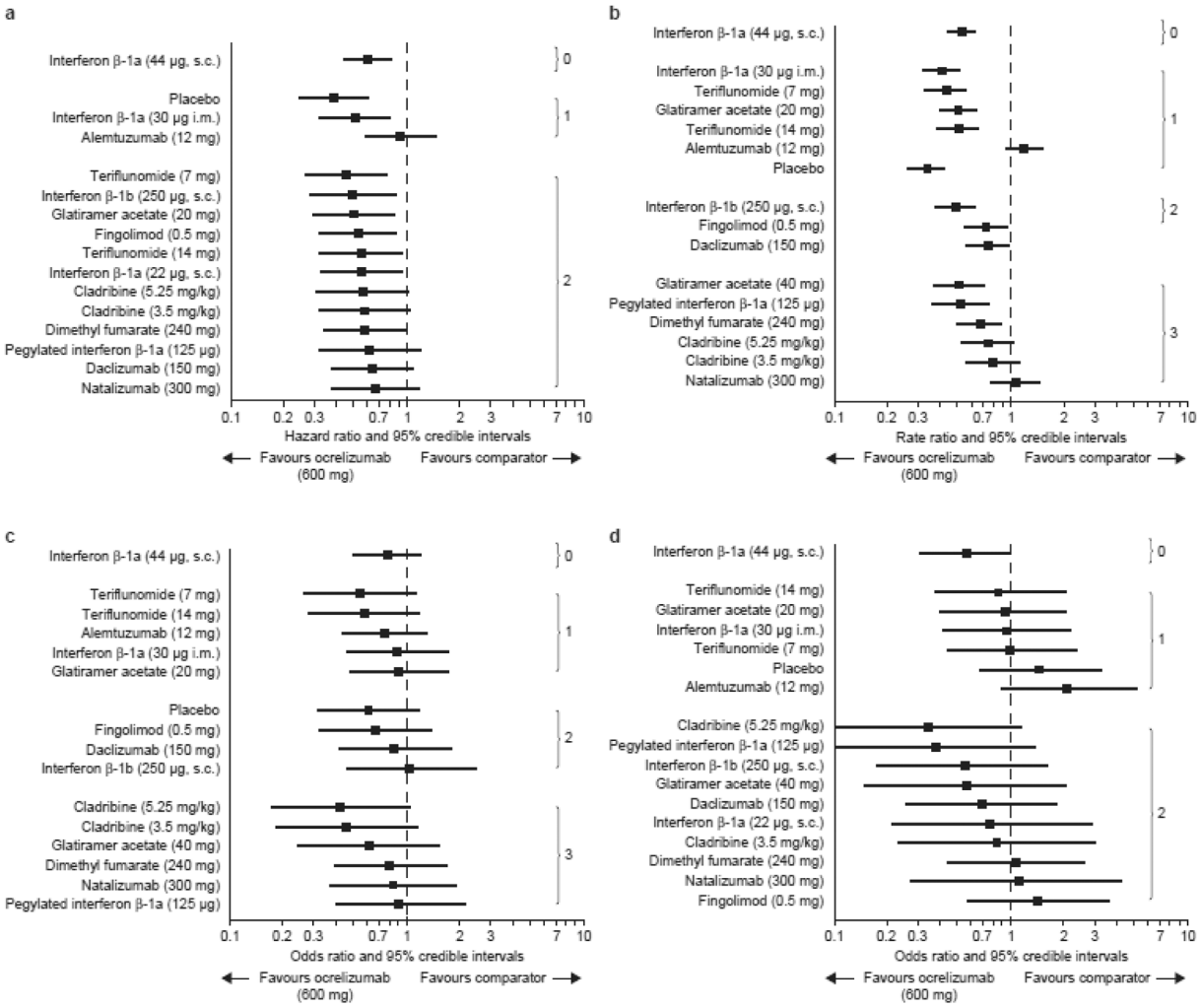

Fig. 2. Base-case forest plots for ocrelizumab versus other DMTs[a] for 12-week CDP (panel a), ARR (panel b), SAEs (panel c) and discontinuation due to AEs (panel d). [a]For treatments for which data were available. The numbers next to the brackets on the right of each panel represent the jumps in the network between ocrelizumab and the other treatments. AE, adverse event; ARR, annualized relapse rate; CDP, confirmed disability progression; DMT, disease-modifying therapy; i.m., intramuscular; SAE, serious adverse event; s.c., subcutaneous.

3.1.4. Discontinuation due to adverse events

The base-case network for discontinuation due to adverse events included 18 different treatments (including placebo) from 31 trials (**Table D17**). Comparisons of ocrelizumab versus other DMTs showed that there was no evidence of a difference between ocrelizumab and any other treatments in the discontinuation of treatment due to AEs (Fig. 2**, panel d**).

3.1.5. Sensitivity analyses and meta-regressions

The results of the base-case NMAs of the key efficacy and safety outcomes were robust to sensitivity analyses and meta-regressions (**Table D19**). The DIC values for sensitivity analyses 1 and 2 were not different enough (i.e. more than ± 3) from the base-case DIC values to be considered meaningful. For each of these outcomes, the DIC values suggest that meta-regression did not improve the model fit over the base-case model (**Table D19**), indicating follow-up time did not modify the relative treatment effects from the base-case NMA, or at least that these models could not demonstrate such an effect (Welton et al., 2012).

3.1.6. Heterogeneity and inconsistency assessments

Most pairwise comparisons had heterogeneity which "might not be important" (I^2 statistics less than or equal to 40%) (Higgins et al., 2011). The consistency models had lower or similar DIC values compared with the inconsistency models, indicating no evidence of inconsistency in the networks (**Appendix B3.3.3; Table D19**).

3.2. Surface under the cumulative ranking curve (SUCRA) radar plot

The SUCRA value rankings reflect the above results and are reported in Table 1. When the probabilities for the two key efficacy outcomes were plotted alongside the two key safety outcomes to provide an overview of all available comparative evidence, ocrelizumab demonstrated a consistently high probability of being ranked as the most effective or tolerable treatment across all four outcomes (Fig. 3).

3.3. Other outcomes analyzed

The base-case network for 24-week CDP included 15 different treatments, including placebo, from 21 trials (**Table D7**), excluding the INCOMIN trial because the 24-week CDP result is regarded as an outlier by MS experts and neurologists (Vartanian, 2003). Comparisons of ocrelizumab versus other DMTs provided evidence that ocrelizumab was more effective in reducing the risk of 24-week CDP than placebo, interferon β-1a 44 µg (Rebif) and teriflunomide 7 mg (**Appendix B3.1; Fig. D3, panel a; Table D15**). The base-case network for the

Table 1
Treatment ranking tables ordered by SUCRA values for key efficacy and safety outcomes.[a]

Annualized relapse rate		12-week confirmed disability progression		Serious adverse events		Discontinuation due to adverse events	
Treatment ranking	%	Treatment ranking	%	Treatment ranking	%	Treatment ranking	%
Alemtuzumab (12 mg)	98.0	Ocrelizumab (600 mg)	95.5	Interferon β-1b (250 μg, s.c.)	81.3	Alemtuzumab (12 mg)[a]	93.9
Natalizumab (300 mg)	92.7	Alemtuzumab (12 mg)	91.9	Ocrelizumab (600 mg)	78.7	Placebo	86.6
Ocrelizumab (600 mg)	88.9	Natalizumab (300 mg)	71.2	Glatiramer acetate (20 mg)	72.9	Fingolimod (0.5 mg)	82.0
Cladribine (3.5 mg/kg)	75.3	Daclizumab (150 mg)	66.0	Pegylated interferon β-1a (125 μg)	71.4	Dimethyl fumarate (240 mg)	64.4
Daclizumab (150 mg)	70.3	Pegylated interferon β-1a (125 μg)	59.6	Interferon β-1a (30 μg i.m.)	71.1	Natalizumab (300 mg)	63.9
Cladribine (5.25 mg/kg)	69.8	Interferon β-1a (44 μg, s.c.)	58.1	Daclizumab (150 mg)	66.1	Teriflunomide (7 mg)	60.6
Fingolimod (0.5 mg)	67.8	Dimethyl fumarate (240 mg)	53.9	Natalizumab (300 mg)	64.2	Ocrelizumab (600 mg)	60.0
Dimethyl fumarate (240 mg)	60.4	Cladribine (3.5 mg/kg)	52.5	Dimethyl fumarate (240 mg)	60.6	Interferon β-1a (30 μg i.m.)	57.3
Interferon β-1a (44 μg, s.c.)	40.3	Cladribine (5.25 mg/kg)	49.1	Interferon β-1a (44 μg, s.c.)	55.6	Glatiramer acetate (20 mg)	54.8
Pegylated interferon β-1a (125 μg)	36.6	Teriflunomide (14 mg)	47.6	Alemtuzumab (12 mg)	51.1	Cladribine (3.5 mg/kg)[a]	49.3
Glatiramer acetate (40 mg)	34.2	Interferon β-1a (22 μg, s.c.)	47.0	Fingolimod (0.5 mg)	40.4	Teriflunomide (14 mg)	48.3
Teriflunomide (14 mg)	34.2	Fingolimod (0.5 mg)	40.5	Glatiramer acetate (40 mg)	35.6	Interferon β-1a (22 μg, s.c.)	44.4
Glatiramer acetate (20 mg)	33.2	Interferon β-1a (30 μg i.m.)	33.2	Placebo	29.1	Daclizumab (150 mg)	35.3
Interferon β-1b (250 μg, s.c.)	26.8	Glatiramer acetate (20 mg)	32.0	Teriflunomide (14 mg)	26.2	Glatiramer acetate (40 mg)	29.2
Teriflunomide (7 mg)	13.0	Interferon β-1b (250 μg, s.c.)	31.0	Teriflunomide (7 mg)	20.9	Interferon β-1b (250 μg, s.c.)	25.2
Interferon β-1a (30 μg i.m.)	8.4	Teriflunomide (7 mg)	18.8	Cladribine (3.5 mg/kg)	14.7	Interferon β-1a (44 μg, s.c.)	22.2
Placebo	0.0	Placebo	2.2	Cladribine (5.25 mg/kg)	10.0	Pegylated interferon β-1a (125 μg)	13.5
						Cladribine (5.25 mg/kg)[a]	9.1

CDP, confirmed disability progression; i.m., intramuscular; s.c., subcutaneous; SUCRA, surface under the cumulative ranking curve.

[a] Because SUCRA values take into account the probability of a treatment being ranked in each position, a treatment with high probabilities of being second, third etc. could still have a high SUCRA value.

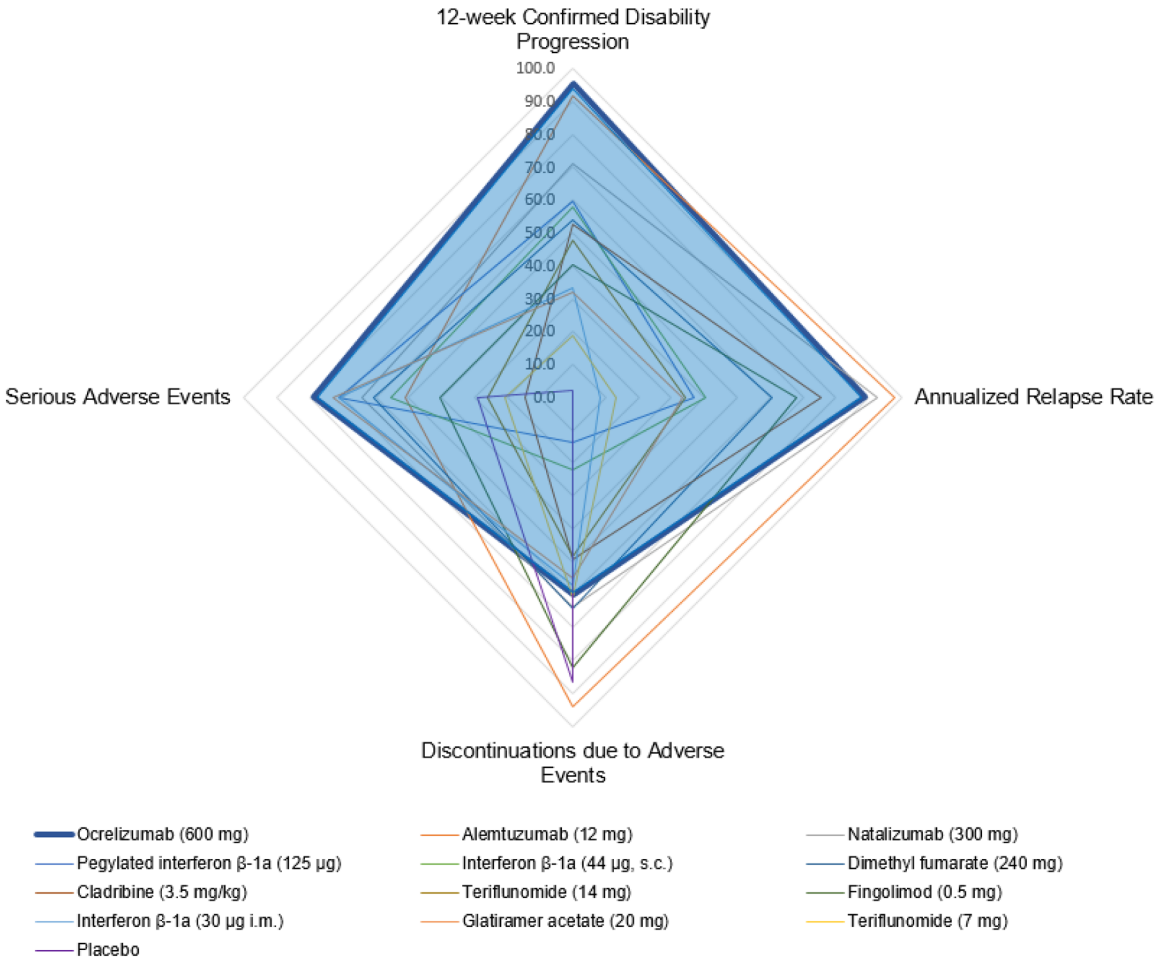

Fig. 3. Radar plot of treatment rankings based on SUCRA values for key efficacy and safety outcomes.[a] [a]For each outcome, treatments towards the outer edge of the plot have SUCRA probabilities closer to 100% and are therefore more effective or more tolerable relative to the other treatments considered. Only treatments which reported on all four outcomes are included in this radar plot (daclizumab was excluded as it was no longer marketed at the time of publication). i.m., intramuscular; s.c., subcutaneous; SUCRA, surface under the cumulative ranking curve.

proportion relapse-free included 18 different treatments, including placebo, from 31 trials (**Table D7**). Comparisons of ocrelizumab versus other DMTs provided evidence that patients receiving ocrelizumab were more likely to remain relapse-free than 14 of the treatments, including placebo (**Appendix B3.1; Fig. D3, panel b; Table D15**). For all-cause discontinuation of treatment, the comparison of 17 treatments in 26 trials (**Table D17**) provided evidence to suggest that patients who receive ocrelizumab are less likely to discontinue than patients who receive pegylated interferon β-1a or interferon β-1a 44 μg (Rebif), but are more likely to discontinue than patients who receive alemtuzumab or natalizumab (**Appendix B3.1; Fig. D3, panel c**).

Most pairwise comparisons had heterogeneity which "might not be important" (Higgins et al., 2011), there was no evidence of inconsistency, and the base-case NMAs were robust to sensitivity analyses (**Appendix B3.2; Table D19**). The SUCRA value rankings for 24-week CDP, proportion relapse-free and all-cause discontinuation of treatment are presented in **Appendix B.3.4** and **Table D21**.

4. Discussion

Ocrelizumab has proven efficacy versus interferon β-1a 44 μg (Rebif) in the pivotal trials OPERA I and OPERA II (Hauser et al., 2017). However, direct comparisons of ocrelizumab with all other DMTs have not been conducted. Therefore, we estimated relative treatment effects for ocrelizumab versus all DMTs currently approved for the treatment of RMS, by examining four efficacy and three safety outcomes in NMAs. These models and the SUCRA ranking values calculated from them suggest that ocrelizumab has an efficacy and safety profile that is superior to or comparable with other available DMTs across all outcomes, except natalizumab and alemtuzumab for one safety outcome: all-cause discontinuation. When interpreting these all-cause discontinuation results, it is necessary to consider possible bias from the inability to discontinue induction treatments once the induction phase has been completed, and the fact that patients were not blinded to the treatment in some pivotal open-label trials. Ocrelizumab provides a treatment option that demonstrates consistent efficacy and safety across all analyzed outcomes (Figs. 2, 3 and **D3**).

The 12-week CDP model provided evidence that ocrelizumab was more effective than 10 treatments, while the 24-week CDP model provided evidence that ocrelizumab was more effective than placebo, interferon β-1a 44 μg (Rebif) and teriflunomide 7 mg, based on the 95% credible intervals. When interpreting these results, it should be considered that the 24-week CDP network has less power to detect a difference between treatments, evidenced by the wider credible intervals for 24-week CDP (**Fig. D3, panel a**) than 12-week CDP (Fig. 2**, panel a**). This is attributable to two things. First, 24-week CDP is, by definition, a rarer event than 12-week CDP. Second, the interconnectedness is greater and there are fewer jumps between ocrelizumab and the other treatments in the 12-week CDP, which decreases the uncertainty in the NMA (Lu, 2004) (**Fig. D1**).

Several NMAs have been performed to compare the treatment effects of DMTs in MS (Fogarty et al., 2016; Huisman et al., 2017; Institute for Clinical and Economic Review, 2017; Tolley et al., 2015; Tramacere et al., 2015; Tsivgoulis et al., 2015; Siddiqui et al., 2017; Melendez-Torres et al., 2017). Most recently in November 2017, Siddiqui et al. compared efficacy across five outcomes: ARR, 12-week CDP, 24-week CDP, proportion relapse-free, no evidence of disease activity (NEDA) and a single combined safety outcome (any AEs) (Siddiqui et al., 2017). These analyses were performed in active relapsing-remitting MS (RRMS) and a subgroup with high disease activity. This manuscript made different analysis assumptions, including the inclusion of trials with a shorter duration in the ARR network but the exclusion of trials under 24 months from the CDP NMAs; they also ran an NMA on the NEDA endpoint despite differences in imaging definitions and monitoring timepoints across trials. In addition, the treatment rankings in Siddiqui et al. were based on the point estimates of treatment effect in the NMAs, which, unlike the SUCRA approach reported here, do not consider the uncertainty of these ranking estimates.

In March 2017, the Institute for Clinical and Economic Review (ICER) report compared efficacy across two outcomes: CDP, which was analyzed with 12-week and 24-week CDP combined, and ARR (Institute for Clinical and Economic Review, 2017). Although results were similar to the analysis reported here, there were differences in methodology. First, ICER did not analyze the proportion relapse-free, and only compared absolute safety outcomes; despite ICER's emphasis of the benefit-risk profiles, ratings seem biased towards efficacy and analysis of safety was limited. Second, the ICER report used event counts for the combined CDP network, whereas this analysis used all available data by combining hazard ratio and count data (Watkins, 2018). Finally, the ICER report included rituximab, used off-label for treatment of RMS, with no approved dose (which was underpowered compared with other treatments in the ARR network), and excluded cladribine, which has since been approved by the EMA.

Fogarty et al. also conducted an NMA for the DMTs approved for the treatment of RRMS up to 2016, and found similar results to the analyses reported here (Fogarty et al., 2016). However, ocrelizumab was not included, because it had not been approved at the time of analysis, and safety outcomes were not considered. Similarly, Tramacere et al. (2015) carried out an NMA for RRMS and analyzed 12-week CDP and 12- and 24-month ARR. Although they considered ARR at two different times, they analyzed only a single safety outcome: withdrawal due to AEs. Neither of these publications analyzed the proportion relapse-free, meaning this analysis is the only peer-reviewed NMA to consider both this outcome and extensive safety outcomes, while including all currently approved DMTs for RMS based on a recent literature review (July 2017).

5. Strengths and limitations

The large number of RCTs reporting data for RMS treatments allowed NMAs assessing all currently approved DMTs. Analyses were performed following the NICE DSU TSDs (Dias et al., 2011b), which provide guidelines for many of the statistical methods used for this analysis. Assessments were made to evaluate whether risk of bias in trials and heterogeneity between trials for each outcome was acceptable (Dias et al., 2011b). SUCRA values were generated alongside the NMAs and provided a more complete overview of all relative comparative evidence. These strengths allowed robust comparisons that can be used to inform reimbursement and treatment decisions.

NMAs of safety outcomes are limited by several factors, including: (i) no trials were statistically powered to analyze safety; (ii) events included were only those recorded during the trial period; and (iii) it is unclear how an induction treatment can be discontinued, which impacts the discontinuation due to AEs and all-cause discontinuation networks. There were several further limitations of this analysis shared by previously published NMAs. First, the validity of these analyses depends on the appropriateness of assuming similarity across trials (e.g. variation in definitions, follow-up time and baseline characteristics). Trial heterogeneity was assessed with sensitivity analyses and meta-regression, which generally agreed with the base-case results. Second, some results were uncertain owing to the limited data available in the networks, and there is risk of bias (the impact of which may have been large if the comparison was central to the network) due to trials that were not double-blinded or reported unexpected dropouts, and missing or inappropriate ITT analyses. Third, the short-term results obtained from RCTs analyzed in this report may not be relevant to longer-term outcomes (i.e. beyond 24 months). Finally, the trials included in all NMAs have been conducted over three decades (since 1987), during which time the natural history of MS has changed; comparing contemporary trials with older trials is likely to result in differential background rates of relapse. In addition, this has led to networks centred around different comparator nodes (e.g. placebo and

interferons) and relying on multiple trials to connect treatments of interest (Fig. 2; **Fig. D1; Fig. D3**).

6. Conclusion

The favourable results of OPERA I and OPERA II demonstrated the efficacy and safety of ocrelizumab in a direct comparison with interferon β-1a 44 µg (Rebif 44 mg). In the absence of head-to-head trials against other DMTs for the treatment of patients with RMS, this NMA provides important evidence on the relative efficacy and safety of ocrelizumab compared with other approved treatments for RMS. This evidence suggests that ocrelizumab has an efficacy superior to or comparable with all other currently approved DMTs across all endpoints analyzed, and a similar safety profile, indicating that it offers a complete and valuable package for the treatment of patients with RMS.

Conflicts of interest

Rachael McCool is an employee of York Health Economics Consortium (YHEC).

Katy Wilson is an employee of York Health Economics Consortium (YHEC).

Mick Arber is an employee of York Health Economics Consortium (YHEC).

Kelly Fleetwood was an employee of Quantics Biostatistics at the time of analysis.

Sydney Toupin is an employee of Quantics Biostatistics.

Howard Thom has received consulting fees from F. Hoffmann-La Roche Ltd., Novartis Pharma AG, and Pfizer Inc.

Iain Bennett is an employee and shareholder of F. Hoffman-La Roche Ltd.

Susan Edwards is an employee and shareholder of F. Hoffmann-La Roche Ltd.

Funding/acknowledgements

Analysis for this work was performed by F. Hoffmann-La Roche Ltd., Quantics Biostatistics and York Health Economics Consortium (YHEC) and funded by F. Hoffmann-La Roche Ltd.

Medical writing support was provided by Fraser Harris of Oxford PharmaGenesis, UK, with funding from F. Hoffmann-La Roche Ltd.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.msard.2018.12.040.

References

Dias, S., Welton, N.J., Sutton, A.J., Ades, A.E., 2011a. NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials. NICE Decision Support Unit. https://www.ncbi.nlm.nih.gov/books/NBK310366/pdf/Bookshelf_NBK310366.pdf (accessed November 2018).

Dias, S., Welton, N.J., Sutton, A.J., et al., 2011b. NICE DSU Technical Support Document 4: inconsistency in networks of evidence based on randomised controlled trials. NICE Decision Support Unit. https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0088912/pdf/PubMedHealth_PMH0088912.pdf (accessed November 2018).

Etemadifar, M., Janghorbani, M., Shaygannejad, V., 2006. Comparison of Betaferon, Avonex, and Rebif in treatment of relapsing-remitting multiple sclerosis. Acta. Neurol. Scand. 113, 283–287.

European Medicines Agency (EMA), 2018. Ocrevus. summary of product characteristics. https://www.ema.europa.eu/documents/product-information/ocrevus-epar-product-information_en.pdf (accessed November 2018).

Fogarty, E., Schmitz, S., Tubridy, N., Walsh, C., Barry, M., 2016. Comparative efficacy of disease-modifying therapies for patients with relapsing remitting multiple sclerosis: systematic review and network meta-analysis. Mult. Scler. Relat. Disord. 9, 23–30.

Gajofatto, A., 2015. Benedetti MD. Treatment strategies for multiple sclerosis: when to start, when to change, when to stop? World J. Clin. Cases 3, 545–555.

Hauser, S.L., Bar-Or, A., Comi, G., et al., 2017. Ocrelizumab versus interferon beta-1a in relapsing multiple sclerosis. N. Engl. J. Med. 376, 221–234.

Higgins, J.P.T., Altmann, D.G., Sterne, J.A.C., 2011. Assessing risk of bias in included studies. In: Higgins, JPT, Green, S (Eds.), Cochrane Handbook for Systematic Reviews of Interventions, . https://handbook-5-1.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm (accessed November 2018).

Hoffmann-La Roche Ltd. Data on File. October 2018.

Huisman, E., Papadimitropoulou, K., Jarrett, J., et al., 2017. Systematic literature review and network meta-analysis in highly active relapsing-remitting multiple sclerosis and rapidly evolving severe multiple sclerosis. BMJ Open 7, e013430.

Institute for Clinical and Economic Review, 2017. Disease-Modifying Therapies for Relapsing-Remitting and Primary-Progressive Multiple Sclerosis: Effectiveness and Value. Final Evidence Report. https://icer-review.org/wp-content/uploads/2016/08/CTAF_MS_Final_Report_030617.pdf (accessed November 2018).

Lavery, A.M., Verhey, L.H., Waldman, A.T., 2014. Outcome measures in relapsing-remitting multiple sclerosis: capturing disability and disease progression in clinical trials. Mult. Scler. Int. 2014, 262350.

Li, T., Puhan, M.A., Vedula, S.S., Singh, S., Dickersin, K., 2011. Network meta-analysis-highly attractive but more methodological research is needed. BMC Med. 9, 79.

Lu, G., 2004. Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat. Med. 23, 3105–3124.

Melendez-Torres, G.J., Auguste, P., Armoiry, X., et al., 2017. Clinical effectiveness and cost-effectiveness of beta-interferon and glatiramer acetate for treating multiple sclerosis: systematic review and economic evaluation. Health Technol. Assess. 21, 1–352.

Merkel, B., Butzkueven, H., Traboulsee, A.L., Havrdova, E., Kalincik, T., 2017. Timing of high-efficacy therapy in relapsing-remitting multiple sclerosis: a systematic review. Autoimmun Rev. 16, 658–665.

National Institute for Health and Care Excellence (NICE), 2015. Single technology appraisal (STA). Specification for manufacturer/sponsor submission of evidence. https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisals/specification-for-company-submission-of-evidence-2015-version.docx (accessed November 2018).

Plummer, M., 2018. JAGS Version 3.3.0 User Manual. http://www.stat.cmu.edu/~brian/463-663/week10/articles,%20manuals/jags_user_manual.pdf (accessed November 2018).

R Development Core Team, 2018. R: a language and environment for statistical computing. https://www.r-project.org/ (accessed November 2018).

Rotstein, D.L., Healy, B.C., Malik, M.T., Chitnis, T., Weiner, H.L., 2015. Evaluation of no evidence of disease activity in a 7-year longitudinal multiple sclerosis cohort. JAMA Neurol. 72, 152–158.

Salanti, G., Ades, A.E., Ioannidis, J.P., 2011. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. J. Clin. Epidemiol. 64, 163–171.

Siddiqui, M.K., Khurana, I.S., Budhia, S., et al., 2017. Systematic literature review and network meta-analysis of cladribine tablets versus alternative disease-modifying treatments for relapsing-remitting multiple sclerosis. Curr. Med. Res. Opin. 1–11.

Su, Y.-S., Yajima, M., 2018. R2jags: Using R to Run 'JAGS'. http://CRAN.R-project.org/package=R2jags (accessed November 2018).

Tolley, K., Hutchinson, M., You, X., et al., 2015. A network meta-analysis of efficacy and evaluation of safety of subcutaneous pegylated interferon beta-1a versus other injectable therapies for the treatment of relapsing-remitting multiple sclerosis. PLoS One 10, e0127960.

Torkildsen, O., Myhr, K.M., Bo, L., 2016. Disease-modifying treatments for multiple sclerosis—a review of approved medications. Eur. J. Neurol. 23 (Suppl 1), 18–27.

Tramacere, I., Del Giovane, C., Salanti, G., D'Amico, R., Filippini, G., 2015. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. Cochrane Database Syst. Rev. 9, CD011381. https://doi.org/10.1002/14651858.CD011381.pub2.

Tsivgoulis, G., Katsanos, A.H., Grigoriadis, N., et al., 2015. The effect of disease modifying therapies on disease progression in patients with relapsing-remitting multiple sclerosis: a systematic review and meta-analysis. PLoS One 10, e0144538.

Turner, R.M., Jackson, D., Wei, Y., Thompson, S.G., Higgins, J.P., 2015. Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis. Stat. Med. 34, 984–998.

Vartanian, T., 2003. An examination of the results of the EVIDENCE, INCOMIN, and phase III studies of interferon beta products in the treatment of multiple sclerosis. Clin. Ther. 25, 105–118.

Watkins, C., 2018. Bennett I. A simple method for combining binomial counts or proportions with hazard ratios for evidence synthesis of time-to-event data. Res. Synth. Methods. https://doi.org/10.1002/jrsm.1301. Epub ahead of print.

Welton, N.J., Sutton, A.J., Cooper, N., Abrams, K.R., Ades, A.E., 2012. Evidence Synthesis for Decision Making in Healthcare. Wiley. John Wiley & Sons Ltd., Chichester, United Kingdom.